

Universiteit Utrecht



Department
of
Mathematics

Repairing near-singularity for dense EMC problems
by adaptive basis techniques
by
Menno E. Verbeek

Preprint
nr. 1125
December 1999

Repairing near-singularity for dense EMC problems by adaptive basis techniques

Menno E. Verbeek*

December 15, 1999

Abstract

A standard boundary element method (BEM) discretisation for the electromagnetic compatibility (EMC) problem is analysed, revealing the near-singularity of the dense linear system for the lower frequency range. Next, a new basis transformation is presented, that separates the dominating capacitive effects from the small, but important inductive effects. On this basis a sparse approximate inverse is constructed, that is used not only as a preconditioner, but also as a smoother for a multi-level preconditioner. The number of matrix-vector multiplications for the iterative solver is further reduced by reusing information: Ritz vectors, obtained from previous right-hand side solves, are injected into the search space of GMRESR, thereby skipping the stagnation phase. Reuse of information from the previous frequency computations has also been attempted, but this is less effective. Test results for the different methods are presented and compared.

1 Introduction

All electrical appliances have to conform the electromagnetic compatibility (EMC) laws, which means that the radiated electromagnetic field must comply to some restrictions. To shorten the design process, it is important to be able to get fast and accurate predictions for EMC properties. This involves calculating the response of a conducting shape to several excitation sources for a wide range of frequencies. These results can also be used for the analysis of an electrostatic discharge. Also, the same type of problem has to be solved for electromagnetic scattering and antenna simulations.

These problems can be solved by solving the Maxwell equations on the domain, using a finite element method (FEM) [8]. However, we are interested in the properties of the radiated field, which theoretically extends to infinity. In practice the domain can be truncated at some distance, but we will have to extend the domain far beyond the size of the object, and impose some artificial boundary condition.

An alternative approach uses a boundary integral formulation, leading to a boundary element method (BEM). This electromagnetic problem can be written as an integral equation on the surface of the conductor. In this way we can calculate the currents resulting from the excitation, which we can use to calculate the radiated fields. Hereby, we reduce the domain of the problem from 3- to 2-dimensional, and skip the need to extend the domain beyond the size of the object. This can dramatically reduce the number of unknowns, but it makes it necessary to solve a dense linear system instead of the sparse system involved with the FE method. The BE method is most effective for problems with lots of empty space.

It is also possible to combine the two methods. If some regions of the domain are dense with conductors, these can be treated with the FE method, while other, more empty regions, including

*Mathematical Institute, Utrecht University, Utrecht, The Netherlands and Philips Research, Eindhoven, The Netherlands. E-mail: verbeek@math.uu.nl

the exterior, are treated with the BE method. This requires some connection mechanism for the different types of regions.

We will study the pure BE method to solve EMC problems. However, the methods described here can also be used for other problems of the same type, and may also lead to good preconditioners for the BEM part of a mixed method.

Modelling with the boundary element method leads to the necessity to solve many large dense linear systems. Because of the bad condition and the density of the matrix this is often done with direct solvers, but we will show that, after constructing a clever basis, an iterative solver combined with a good preconditioner may lead to a faster method.

2 Physics and discretisation

The problem of interest is written as a boundary integral equation on the conductor surface Γ [1]

$$\frac{i}{\omega} \int_{\Gamma} \nabla_{\Gamma} G(\mathbf{x}, \mathbf{x}') \nabla'_{\Gamma} \cdot \mathbf{J}(\mathbf{x}') d^2 x' + \frac{i\omega}{c^2} \int_{\Gamma} G(\mathbf{x}, \mathbf{x}') \mathbf{J}(\mathbf{x}') d^2 x' + Z(\mathbf{x}) \mathbf{J}(\mathbf{x}) = \mathbf{E}_{\Gamma}^E(\mathbf{x}) \quad \forall \mathbf{x} \in \Gamma \quad (1)$$

Here, G is the Green function

$$G(\mathbf{x}, \mathbf{x}') = \frac{e^{-ik|\mathbf{x}-\mathbf{x}'|}}{|\mathbf{x}-\mathbf{x}'|} \quad , \quad (2)$$

ω is the angular frequency of the excitation, $k = \omega/c = 2\pi/\lambda$ is the corresponding wave number and ∇_{Γ} indicates the gradient or divergence restricted to the conductor surface Γ . The first integral is called the capacitive part and is due to the surface charge density ρ :

$$\nabla_{\Gamma} \cdot \mathbf{J} + i\omega\rho = 0 \quad . \quad (3)$$

The second integral is called the inductive part and is due to the surface current density \mathbf{J} . The third term is due to the surface impedance Z which is often very small. The right-hand side is the external electric field, which acts as a the driving force. Equation (1) is called the Electric Field Integral Equation (EFIE). The EFIE is known to lead to problems with internal resonances [1], but since our models only consist of thin plates and wires, this should not be a problem here.

To solve the linear EFIE (1), we discretise it with a Galerkin approach, which is also called the method of moments in this context. We rewrite the EFIE in a weak form, where we use only a finite number of test functions $\Psi_i(\mathbf{x})$, $i = 1 \dots n$, and we use the same functions as a basis for the current density :

$$\mathbf{J}(\mathbf{x}) = \sum_i x_i \Psi_i(\mathbf{x}) \quad . \quad (4)$$

This results in a set of linear algebraic equations that can be written in matrix form

$$Ax = b \quad , \quad (5)$$

with

$$\begin{aligned} A &= C + L + R \\ C_{ij} &= \frac{i}{\omega} \int_{\Gamma} \nabla_{\Gamma} \cdot \Psi_i(\mathbf{x})^* G(\mathbf{x}, \mathbf{x}') \nabla'_{\Gamma} \cdot \Psi_j(\mathbf{x}') d^2 x' d^2 x \\ L_{ij} &= \frac{i\omega}{c^2} \int_{\Gamma} \Psi_i(\mathbf{x})^* G(\mathbf{x}, \mathbf{x}') \Psi_j(\mathbf{x}') d^2 x' d^2 x \\ R_{ij} &= \int_{\Gamma} \Psi_i(\mathbf{x})^* Z(\mathbf{x}) \Psi_j(\mathbf{x}) d^2 x \\ b_i &= \int_{\Gamma} \Psi_i(\mathbf{x})^* \mathbf{E}^E(\mathbf{x}) d^2 x \quad . \end{aligned} \quad (6)$$

The conductors are composed of thin wires and plates. The conductor surface Γ should therefore contain a front and back plane for each plate. If the plate is thin enough, we can combine the current on the front and back surface in one current. This reduces Γ to one plane per plate and helps to significantly reduce the numerical work. For the thin wires we will use only currents along the wire, independent of the angle.

3 Matrix properties

We can derive some properties of the matrix A by examining the integrals in (6). First we note that the matrix A is a dense, complex symmetric matrix. Next we see that A has the largest elements for close interactions, which are mostly imaginary. This means that A/i is relatively close to a real symmetric (indefinite) matrix. Consequently, most eigenvalues of A will be close to the imaginary axis. Using energy conservation arguments, we can show that all eigenvalues of A must have positive real parts. However, for most eigenvalues, the real part is small compared to the imaginary part.

Very important for the spectrum of A is the fact that, in general, the capacitive matrix C has a large null space. Cx represents the electric field due to the charge accumulation that results from a current represented by x . If this current is a divergence free current, there will be no charge accumulation and Cx will be zero. In the discretised case, such divergence free currents are the currents that form a closed loop. For each independent loop there will be an independent divergence free current and together they span the null space of C . Depending on the model geometry, its dimension will typically be in the order of $1/2$ of the dimension of C .

From the definitions (6) we see that $C \propto \frac{1}{\omega}$, while $L \propto \omega$. This implies that for small enough frequencies ω (λ much larger than the size of the model), the inductive part L will be much smaller than the capacitive part C . Since the resistive part R is usually very small, the capacitive effects will dominate. However, on the null space of C only the inductive and resistive effects are left, which are relatively small. This results in a cluster of relatively small eigenvalues of the dimension of the null space of C . These eigenvalues can be several orders of magnitude smaller than the remainder of the eigenvalues and this is a big problem for iterative solvers.

To make this more precise, we analysed the continuous spectrum of an infinite board and infinite wire in appendices A.1 and A.2. When using an approximately uniform grid, this continuous spectrum can give an indication of the behaviour of the discrete spectrum of A . In the case of a discretised board/wire of finite size, we can expect the eigenvectors to correspond approximately to modes that fit on the board/wire and on the discretisation. This gives upper and lower cutoff frequencies of the order of $2\pi/L \leq |q| \leq 2\pi/h$, with uniformly distributed values in between. Since we may expect the solution to oscillate with wavelength λ over interior regions, we have to choose a grid sufficiently fine to be able to accurately represent this. In practice $h \leq \lambda/20$ is used, which means that $2\pi/h \geq 40\pi/\lambda = 20k$. The result is that most of the eigenvectors of the matrix will come from the region beyond the resonance ($|q| > k$).

For the inductive part of the spectrum, the eigenvalues are proportional to $1/|q|$ so that the high frequency eigenvalues tend to cluster towards the cutoff eigenvalue, which is nice for the convergence of an iterative solver. However, this cutoff eigenvalue will tend to zero for increasingly fine grids, leading to a growing condition number of the matrix. The capacitive part of the spectrum is proportional to $|q|$, which does not lead to clustering and is thus not favourable for the convergence of an iterative solver. Here, finer grids also lead to a growing condition number.

In total this gives us the following estimate for the (spectral) condition number of A for low frequencies and no resistance

$$\kappa(A) \gtrsim \frac{4\pi^2}{h^2 k^2} = \frac{\lambda^2}{h^2} \quad , \quad (7)$$

where possible resonances may lead to even worse values. Note that for low frequencies the edge effects and the geometry dictate the element size h , leading to a much smaller h than required by

the inequality used above. As a conservative example, for a pretty coarse grid size of 1 cm and an intermediate frequency of 10 MHz ($\lambda=30$ m) we still get $\kappa(A) \gtrsim 10^7$.

For the higher frequencies, not the high frequency modes but the resonances lead to problems. Since the resonances depend strongly on the geometry and the frequency, this condition number is quite hard to estimate.

4 Constructing a new basis

For low frequencies, standard preconditioning techniques will encounter severe difficulties because of the large cluster of small eigenvalues dictated by $L + R$. Since the contribution of $L + R$ to the elements of A will be very small compared with the contribution of C , a standard preconditioner will not be able to capture the behaviour of A on the null space of C . It is possible to precondition the part dominated by C , but the large cluster of small eigenvalues will remain unaffected.

To overcome the problem of the cluster of small eigenvalues we will try to separate the contribution of $L + R$, visible in the small eigenvalues, from the contribution of C , seen in the large eigenvalues. To achieve this, we will use a basis for the small eigenvalue subspace and a basis for the remaining large eigenvalue subspace.

From the topology of the discretisation we can construct a set of basis vectors K_l for the null space of C , which contains all the independent loop currents. We then have that $CK_l = 0$, and since $C^T = C$ also $K_l^T C = 0$. Note that in practice K_l will contain mainly small local loops on the boards and only a few global loops, resulting in a sparse K_l . The basis vectors K_c for the remainder of the space can still be chosen freely. K_c need not be orthogonal to K_l , but should not make small angles with K_l either, because this would lead to numerical problems in the solution process. We now define the basis transformation $Q = (K_l, K_c)$ by

$$A_Q \equiv Q^T A Q = \begin{pmatrix} K_l^T(L + R)K_l & K_l^T(L + R)K_c \\ K_c^T(L + R)K_l & K_c^T(C + L + R)K_c \end{pmatrix}. \quad (8)$$

On this new basis, C is restricted to one block. As a result, the contribution of $L + R$, that determines the behaviour of A on the null space of C , is not hidden behind the C any more, and a preconditioner may be able to capture it. Note that A_Q is still symmetric.

The question remains how to select K_c , the second part of the new basis. Since the first part deals with chargeless currents, K_c should represent all possible charges. Furthermore, we know that the operator associated with the potential due to charges is quite well behaved (see appendix A). In electrostatics this is given by

$$V(\mathbf{x}) = \int_{\Gamma} G(\mathbf{x}, \mathbf{x}') \rho(\mathbf{x}') d^2 x', \quad (9)$$

where ρ is the surface charge density and V the electrostatic potential. We will try to get a similar operator for $K_c^T C K_c$, by choosing the right basis vectors K_c . Since capacitive effects dominate this part of A_Q , we hope $K_c^T(C + L + R)K_c$ will behave alike. The basis vectors K_c describe a current, which means that they always have total zero charge. It is therefore impossible to get pure charge basis functions. However, we can devise currents so that one charge element has a unit positive charge and the others have a small negative charge, such that the total charge is again zero. Filling K_c with these vectors makes sure that $K_c^T C K_c$ is (up to a constant) the discretised operator of equation (9).

This can also be expressed in matrix terms. From the EFIE (1) or from equation (3) and $\mathbf{E} = -\nabla V$ we can see that the “current \rightarrow field” matrix C can be composed from the “charge \rightarrow potential” matrix D associated with (9), the surface gradient matrix R and the surface divergence matrix R^T :

$$C = R D R^T. \quad (10)$$

C is the $n \times n$ matrix in (6) and D is an $m \times m$ matrix associated with (9), with n the total number of current degrees of freedom and m the number of charge degrees of freedom. The surface

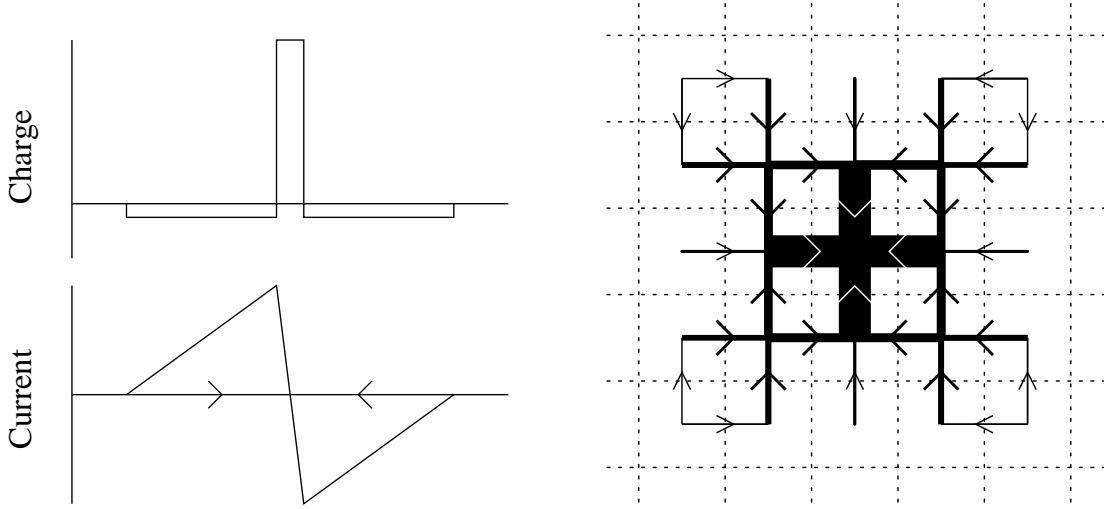


Figure 1: Left: Current and charge of the quasi-charge basis functions on the wires. Right: Current for the quasi-charge basis functions on the boards. The dotted lines denote the grid, the arrows the current (thickness is proportional to the current).

gradient matrix R is a sparse $n \times m$ matrix with on each row one $+1$ and one -1 element. From (10), we see that K_l , the null space of C , is equal to the null space of R^T . This null space K_l has at least size $n - m$, but since R does not have full rank it is larger. For each disconnected part of the conductor surface the gradient of a constant function is zero, thus decreasing the rank of R by one, and increasing the dimension of the null space of C by one. As a result, the dimension of this null space is $n - m + p$, where p is the number of disconnected parts in the conductor surface. We would like $K_c^T C K_c = K_c^T R D R^T K_c$ to have the same nice behaviour as D , so we would like to choose K_c such that $K_c^T R = R^T K_c = I$. Unfortunately this is impossible because R does not have full rank. We can choose K_c such that $K_c^T R = R^T K_c = 1 - E$, where $1 - E$ is a projection on the space of total zero charge per disconnected conductor. In the case of one connected conductor ($p = 1$), the matrix E has all elements equal to $1/m$. We now see that $K_c^T C K_c = K_c^T R D R^T K_c = (1 - E) D (1 - E)$, which represents D restricted to the space of total zero charge and potential. This is physically the same, because the total charge is zero and only potential differences are important.

This ideal K_c can be constructed, but it takes too much time and it will lead to a dense matrix K_c , which makes basis transformations expensive. However, we can make a cheap sparse approximation by quasi-charge currents. These are currents that transport charge from a region around a particular charge element to that central charge element. The resulting charge is large positive on the central charge element and small negative on the remainder of the region. In the case of a wire, the quasi-charge current is fully determined by the required charge distribution. For the board, we have some freedom in choosing the currents and that freedom can be used to try to minimise $K_l^T K_c$. The quasi-charge currents we used are shown in Figure 1.

Choosing these quasi-charge currents for the second part of the new basis will lead to an “over-complete” basis Q with $(n - m + p) + m = n + p$ vectors. The null space of Q will correspond to the null space of K_c and its basis will be formed by the vectors with a constant quasi-charge on the disconnected parts of the conductor. Since we now know these vectors, this subspace can be deflated in the iterative solver.

The new Q basis can also be seen from an other point of view. The operator K_l^T can be interpreted as a discretised version of the surface rotation, while the currents in K_l are divergence free. The currents in K_c have a strongly peaked divergence (charge) and a relatively low rotation. This basis transformation might thus be related to a splitting in divergence free and rotation free

currents. A side effect is that $K_l^T K_c$ is relatively small, which results in small $K_l^T A K_c$ and $K_c^T A K_l$ due to a relatively dominant and constant diagonal of A .

As a result the $K_l^T A K_l$ part is decoupled rather well from the $K_c^T A K_c$ part. This decoupling leads to an effective block-diagonal preconditioning with separate preconditioners for the $K_l^T A K_l$ and the $K_c^T A K_c$ blocks.

5 Preconditioning

We can now use a conventional preconditioner for the matrix A_Q , but we will have to adjust this to the density of A_Q . We propose to use the sparse approximate inverse of Grote and Huckle [3], also because of its potential parallelism. This approach is very expensive for a dense matrix, so we will apply this to a sparsified matrix A_Q^{sp} . We get this sparsified matrix by using only those matrix elements that correspond to interactions between physically close elements. To gain even more, the sparsity pattern of the approximate inverse can be fixed in advance, again by using a small-distance criterion. The results is that we do a simple Frobenius norm minimisation

$$\min_M \|A_Q^{\text{sp}} M - I\|_F \quad (11)$$

where the sparsity pattern of M is fixed. We will refer to this as the sparse approximate inverse (SAI) in the remainder of this paper.

In the sparsification of the matrix, we throw away the terms that are important only for the long range behaviour. Hence we cannot expect a good preconditioner for long range effects. For the highly oscillating eigenfunctions, the resulting long distance field will be negligible because of the cancellation. For these highly oscillating eigenfunctions the near interaction terms are the most important ones and they can be represented quite well in the preconditioner. This results in an effective preconditioner for the removal of the highly oscillating modes corresponding to the extreme eigenvalues that caused the bad condition (7) for low source frequencies. This is precisely what we see in practice, the best preconditioned eigenfunctions are the fast oscillating ones, and the bad ones are only varying over larger distances. Fourier analysis confirms that preconditioning with the exact inverse of the truncated operator works well for highly oscillating functions; for more details see appendix A.3.

A problem with the SAI preconditioner is the convergence behaviour when the discretisation is refined. To maintain convergence speed we need to increase the number of nonzeros per row in A_Q^{sp} and M as h^{-2} , which leads to rapidly growing costs.

If we combine the SAI preconditioner with a preconditioner for the slowly varying eigenfunctions, using a coarse grid, we obtain a multilevel preconditioner [4]. It is easy to construct a coarser basis for the two new scalar quantities and we put these coarse basis functions in a tall rectangular interpolation matrix P , and define the accompanying restriction to be P^T . We form the corresponding coarse grid matrix directly from the fine grid matrix by

$$A_{QP} = P^T A_Q P \quad (12)$$

As a smoother we use the sparse approximate inverse. This gives us all the ingredients for a multigrid V-cycle that approximates the solution of $A_Q x = b$, and can be used as a preconditioner for an iterative solver like GMRES [7]. The V-cycle consists of a recursion of 3 steps :

- **Pre-smoothing:** The high-frequency modes in the residual b are damped by multiplication with the smoother M :

$$\begin{aligned} x_1 &= Mb \\ r_1 &= b - A_Q x_1 = (1 - A_Q M)b \end{aligned} \quad (13)$$

- **Coarse grid correction:** Now that the residual r_1 is smooth, it can be represented rather well on the coarser grid ($r_c = P^T r_1$). The much smaller coarse grid system can be inverted

approximately ($x_c \approx A_{QP}^{-1} r_c$) and the solution can be interpolated to the finer grid for the correction of the previous solution :

$$\begin{aligned} x_2 &= x_1 + P A_{QP}^{-1} P^T r_1 \\ r_2 &= b - A_Q x_2 = (1 - A_Q P A_{QP}^{-1} P^T) r_1 \end{aligned} \quad (14)$$

- **Post-smoothing:** The high frequency modes are again removed from the residual :

$$\begin{aligned} x_3 &= x_2 + M r_2 \\ r_3 &= b - A_Q x_3 = (1 - A_Q M) r_2 \end{aligned} \quad (15)$$

The coarse grid solve will be done by another V-cycle on the coarse grid, introducing an even coarser grid. This will continue until the grid is coarse enough to do a direct solve. For our problem we coarsen the grid by a factor of approximately 2 in each direction leading to around 4 times less coarse grid unknowns. To perform the V-cycle on the coarser grid, we also need a smoother for that level. Here, we again use the SAI preconditioner, now applied to $A_{QP}^{\text{sp}} = P^T A_Q^{\text{sp}} P$.

Note that the calculation of A_Q from A is very expensive because of the large and dense matrix A . However, we do not need to know A_Q explicitly, but we only need a sparsified version for the SAI smoother, which is not too expensive to construct. The matrix vector multiplications with A_Q can be performed by consecutive multiplication with Q , A , and Q^T . A_{QP} is calculated explicitly using $(QP)^T A (QP)$, which is cheaper because of the low column dimension of QP .

Note that this method for obtaining the coarse grid matrix cannot trivially be used in combination with a matrix-free matrix vector multiplication like the fast multipole methods. However, we only need to know A explicitly to be able to do matrix vector multiplications with the coarse grid matrices. If we have an efficient matrix-free way to calculate matrix vector multiplications on all levels, we do not need to know A explicitly. By combining the different levels of interpolations and restrictions with the different levels of projections in the fast multipole method, in principle a matrix-free multigrid method is possible.

6 Reusing information

In practical EMC simulation we have to solve the system (5) for multiple right-hand sides and multiple frequencies. We would like to make optimal use of this by maximally reusing information.

The reuse of the GMRES polynomial for multiple right-hand sides, as done as in most block methods, is not advantageous in our case. In general, this will increase the number of iterations (the fixed GMRES polynomial is optimal only for one right-hand side) and it may only save on the GMRES inner products, vector updates and the reduced system solve, which are not the cpu-dominating factors here. Simoncini and Gallopoulos compared some methods and proposed their own MHGMRES method [9], but they still assume that the matrix vector product is relatively cheap in order to make the method profitable. In our case, the dense matrix vector multiplication will dominate the cpu-costs, and only a reduction in the number of matrix vector products can significantly decrease the work. To accomplish this, we have to reuse matrix vector multiplication information from previous solves with the same matrix. This can be done by starting the next solve with the search space from the previous solve, for which we still know the image under A . For this we can use the GCR method [2] with some additions. What we get is GMRESR [10], which allows the injection of an arbitrary search space and variable preconditioning. In [11], GMRESR is also used to inject a search space from previous right-hand side solves, but we will do it slightly different.

When solving $Ax = b$ with GMRESR, we build a search space V and shadow space U , such that $AV = U$ and $U^H U = 1$, and chose the best approximation to the solution from the search space V . If we start V with b and expand it with the residual in each step, then we are building the same Krylov subspace as in ordinary GMRES and, in exact arithmetic, we get the same results for GMRESR and GMRES. However, since we store the shadow space separately, V is not restricted

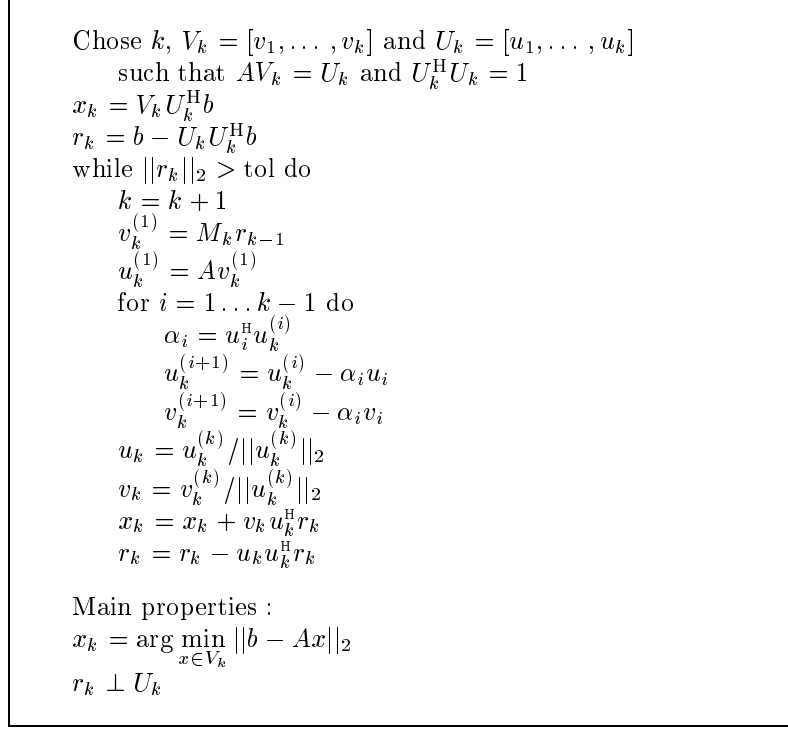


Figure 2: The GMRESR algorithm with search space injection and variable preconditioner M_k .

to Krylov subspaces, but we can use any subspace. This implies that we can expand V differently. For fast convergence, we have to expand V with an approximate solution to $Av = r$, where r is the residual. The way this approximation is made (the preconditioner) can change from step to step, leading to a variable preconditioned method. An other degree of freedom is that we can start the iteration with any pair (V, U) that satisfies $AV = U$ and $U^H U = 1$. We will call this *search space injection*. The GMRESR algorithm with search space injection and variable preconditioning is shown in Figure 2. Unfortunately we have to pay for this flexibility: we have to keep and update both V and U , leading to twice as many vector updates and double the memory requirements compared with GMRES. However, since our matrix-vector multiplication is the most expensive part in terms of cpu-time and memory requirements, this is not a big problem for our application.

The complete injection of the previous V and U will lead to a large search space, and this will lead to much work in the orthogonalisation and it might also lead to memory problems. To prevent this, we want to reuse only the most relevant information. To separate the relevant from the less relevant information, we calculate Ritz values and vectors from V and U , and select those Ritz vectors that represent eigenvalues that cause slow convergence. These Ritz vectors can now be injected in the next solve.

From a pair (V, U) we can calculate harmonic Ritz vectors and values by calculating the eigenvalues and vectors of $U^H V$:

$$\frac{1}{\lambda} y = U^H V y \quad \Leftrightarrow \quad (16a)$$

$$y = \lambda U^H V y \quad \Leftrightarrow \quad (16b)$$

$$U y = \lambda U U^H V y = \lambda V y - \lambda(1 - U U^H) V y \quad \Rightarrow \quad (16c)$$

$$A V y = \lambda V y + r \quad \Rightarrow \quad (16d)$$

$$A x = \lambda x + r \quad . \quad (16e)$$

Note that the residual r is orthogonal to U . In the first step we assume that $U^H V$ is non-singular,

thus $\lambda \neq 0$. If U and V are near orthogonal, the projection in (16c) will move almost all information to the residual r , leading to very bad Ritz vectors and values. In practice we have not observed this, but if it would happen this would only imply that we will not see any acceleration from the injection of these vectors.

Since the preconditioned matrix determines the convergence behaviour, we would like to reuse the Ritz-vectors and -values of the preconditioned system in further solves. We therefore have to use implicit preconditioning, where the A in the algorithm is the preconditioned matrix. As a consequence we can not use a variable preconditioner.

The Ritz vector injection really improves the robustness of the method. It results in the removal of the stagnation phase in the convergence, as is seen in Figure 3 (see the discussion in section 7), and it will make the cpu-time needed to solve the system less dependent on the problem. Only for the first right-hand side we need to overcome the stagnation phase, but the other right-hand sides will require an, almost constant, small number of iterations.

When solving the same problem for different frequencies, it is possible to use the solutions obtained with the previous frequencies as an initial guess for the new frequency. In order to predict the way the solution will change when the frequency is changed, we investigate how the operator changes when the frequency changes. This change is nonlinear and it has two sources. First, the Green function (2) changes, which has a complicated effect. Second, the factor $1/\omega$ in the capacitive part and the factor ω in the inductive part of the EFIE equation (1) give a more traceable change. This second change implies that the capacitive and inductive parts of the solution scale as ω and $1/\omega$, respectively. Since we separated these parts in the basis transformation, we can use this scaling as a crude predictor. This strategy works well for simple problems, low frequencies, and small frequency differences. Since it is very expensive to calculate the new matrix A for a new frequency, it is not an option to go through the frequency domain with very small steps, as is sometimes done in continuation problems. The only way to make this effective is to get a better predictor scheme.

One way of improving the predictor is to exploit a longer frequency history. The “Marching-on-in-Frequency” approach described in [6] has a slightly different background, but the idea can be used and enhanced here. The idea is to use the linear combination of the solutions at previous frequencies that minimises the residual, as a starting point for the new solve. In view of the different behaviour of the capacitive and inductive parts of the solution, this can be optimised by allowing different coefficients for the different components of the previous solutions. Experiments indicate that for very small frequency steps the residual reduction can be considerable, up to only one order of magnitude from the residual of the previous solutions, if 5 or 6 frequencies were used. For the larger frequency steps, we could get up to 2 orders of magnitude reduction using 3 previous frequencies; exploiting a longer history could not decrease the residual further.

For second and later right-hand sides it is possible to combine this strategy with the previously described Ritz vector injection in order to overcome initial stagnation. We now reuse both the solutions of the previous frequencies and the Ritz vectors of the previous right-hand sides (at the same frequencies), which gives good results for small frequency steps (see Figure 3). However, for the first right-hand side we do not have the Ritz vectors. Reusing Ritz vectors from previous frequencies does not work well since the most essential eigenvectors change most rapidly with the frequency. The stagnation of the solver for the first right-hand side is, so far, an unsolved problem.

7 Experimental results

In our experiments we have used 3 different methods for the solution of the linear systems :

- A direct solver from LAPACK,
- Full GMRESR with the sparse approximate inverse (SAI) preconditioning,
- Full GMRESR with the multigrid V-cycle (MG) as preconditioner.

The iterative solvers were terminated when the relative residual norm was less than 10^{-8} .

We have tested the different methods for a test problem with two boards, two connecting wires, and a long antenna wire. The boards are square and their physical size is 2×2 meter, the

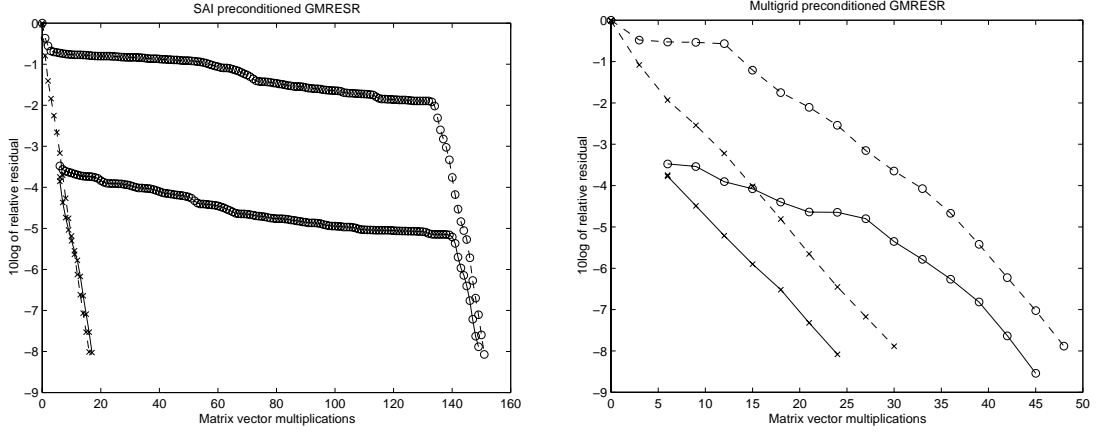


Figure 3: The result of reusing information. All lines for 100MHz, first (o) and second (x) right-hand side, with (—) and without (---) information from 3 previous frequencies (1MHz steps).

n	calc. A	direct	GMRESR & SAI			GMRESR & MG		
320	6	3	5 +	4 =	9 (41)	6 +	2 =	7 (9)
672	28	30	14 +	25 =	39 (62)	18 +	9 =	26 (10)
1430	131	405	29 +	164 =	193 (94)	42 +	52 =	95 (14)
2102	307	1461	47 +	455 =	502 (125)	79 +	120 =	199 (15)
2900	537	3926	73 +	1046 =	1119 (151)	134 +	243 =	377 (16)
measured order	1.7	3.1	1.4	2.6	2.5 (0.6)	1.6	2.2	2.0 (0.2)
expected order	2.0	3.0	1.0	2+a	2+a (a)	2.0	2.0	2.0 (0)

Table 1: Cpu-time in seconds, crude order estimation and expected order for the calculation of A and the 3 solution methods. Times are split in preparation and solve time. The number of iterations is shown in parenthesis.

antenna is about 4 meter long. This is still a relatively simple geometry, but it already gives a good impression of the behaviour of the different methods. Figure 3 shows an example of the convergence history using different combinations of the SAI and MG preconditioners and the information reusing methods described in section 6. In Table 1 we show the cpu-times for one right-hand side and one frequency of 100 MHz ($\lambda = 3$ m), using 5 different grid sizes ranging from 12 to 40 elements per wavelength. Note that since the physics is scalable, the numerical results are the same when all lengths (including the wavelength) are scaled.

We may expect that the effectiveness of our approach is frequency dependent. The idea of the basis transformation (section 4) was inspired by low frequency problems and based on relatively strong capacitive effects. From Figure 4 we see for which frequency range this approach works well. We see that the convergence deteriorates as the frequency increases. For the largest frequencies the solvers did not converge within an acceptable number of iterations. For these frequencies the wavelength λ is so small compared with the element size ($h = 7.4$ cm) that this discretisation does not make much sense. Drastic refinement of the discretisation was not possible because of computer memory limitations. Reducing h with a factor of 2 would result in a 2 Gbyte matrix. This will only be possible on a parallel machine or with a matrix free method for the matrix vector multiplication.

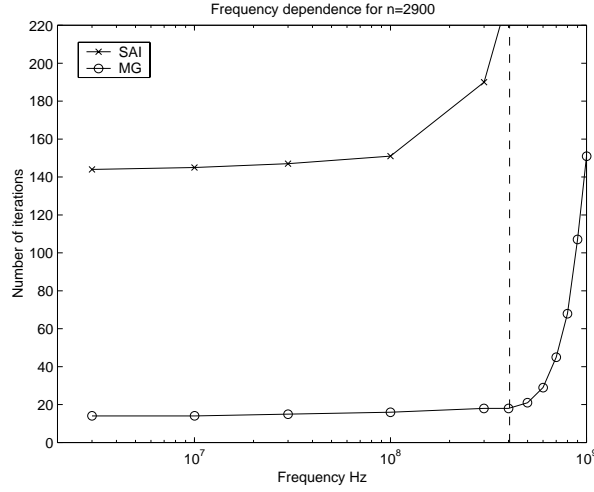


Figure 4: Frequency dependence for problem size $n = 2900$. The maximum element size was $h = 7.4$ cm. The vertical line shows the frequency where $\lambda = 10h$.

8 Conclusion

We have seen that by constructing a better basis, we have effectively changed a linear system that was difficult to solve into one that could be preconditioned with (slightly adjusted) standard methods like SAI and multigrid, and this leads to effective iterative methods. By reusing information we can reduce the work even further. For the dense linear systems in EMC problems, we have shown that, with these techniques, iterative methods can be substantially faster than a direct solver.

The next step will be to get rid of the explicit use of the dense matrix, by using a matrix-free method which hopefully leads to further improvement for the larger problems.

Acknowledgements

I would like to thank Henk van der Vorst (Utrecht University, The Netherlands), Wil Schilders, Jos Bergervoet, and Ronald Rietman (Philips Research, The Netherlands) for posing the problem and for our fruitful discussions.

A Fourier analysis

A.1 The infinite board

In this section we will derive exact eigenfunctions and -values for the continuous infinite board problem. We look at the EFIE (1) where the conductor surface is an infinite plane with constant surface impedance Z . Any current can be decomposed in longitudinal and transversal plane waves so we will look at the result of the action of the operator on the left hand side of equation (1) on the separate plane waves.

We consider general plane waves with wavenumber \mathbf{q} and current direction \mathbf{a}

$$\mathbf{J}_{\mathbf{q}}(\mathbf{x}) = \mathbf{a}e^{i\mathbf{q}\cdot\mathbf{x}}. \quad (17)$$

The inductive field will be given by

$$\begin{aligned}
\mathbf{E}_{\mathbf{q}}^L(\mathbf{x}) &= \frac{ik}{c} \int G(\mathbf{x}' - \mathbf{x}) \mathbf{J}_{\mathbf{q}}(\mathbf{x}') d^2 x' \\
&= \frac{ik\mathbf{a}}{c} \int G(\mathbf{x}' - \mathbf{x}) e^{i\mathbf{q} \cdot \mathbf{x}'} d^2 x' \\
&= \frac{ik\mathbf{a}}{c} e^{i\mathbf{q} \cdot \mathbf{x}} \int G(\mathbf{y}) e^{i\mathbf{q} \cdot \mathbf{y}} d^2 y \\
&= \frac{ik\mathbf{a}}{c} e^{i\mathbf{q} \cdot \mathbf{x}} \hat{G}(\mathbf{q}) ,
\end{aligned} \tag{18}$$

where $G(\mathbf{y}) = \frac{e^{-ik|\mathbf{y}|}}{|\mathbf{y}|}$ is the Green function and $\hat{G}(\mathbf{q})$ is its 2-dimensional Fourier transform. This Fourier transform can not be calculated directly but when we first calculate the Fourier transform of the regularised function $\frac{e^{-ik|\mathbf{y}|}}{|\mathbf{y}|} e^{-\mu|\mathbf{y}|}$ and then let μ go to zero we find that

$$\hat{G}(\mathbf{q}) = -i \frac{2\pi}{\sqrt{k^2 - |\mathbf{q}|^2}^*} \quad \text{for } k \neq |\mathbf{q}| , \tag{19}$$

where the $*$ denotes the complex conjugation of the root. In a similar fashion, we find the capacitive field

$$\begin{aligned}
\mathbf{E}_{\mathbf{q}}^C(\mathbf{x}) &= \frac{i}{ck} \nabla \int G(\mathbf{x}' - \mathbf{x}) \nabla' \cdot \mathbf{J}_{\mathbf{q}}(\mathbf{x}') d^2 x' \\
&= -\frac{i}{ck} (\mathbf{q} \cdot \mathbf{a}) \mathbf{q} e^{i\mathbf{q} \cdot \mathbf{x}} \hat{G}(\mathbf{q})
\end{aligned} \tag{20}$$

and the resistive part

$$\mathbf{E}_{\mathbf{q}}^R(\mathbf{x}) = Z \mathbf{a} e^{i\mathbf{q} \cdot \mathbf{x}} . \tag{21}$$

Adding this together we find

$$\mathbf{E}_{\mathbf{q}}(\mathbf{x}) = \left(\frac{ik}{c} \hat{G}(\mathbf{q}) \mathbf{a} - \frac{i}{ck} (\mathbf{q} \cdot \mathbf{a}) \hat{G}(\mathbf{q}) \mathbf{q} + Z \mathbf{a} \right) e^{i\mathbf{q} \cdot \mathbf{x}} , \tag{22}$$

and we see that $\mathbf{J}_{\mathbf{q}}$ is an eigenfunction ($\mathbf{E}_{\mathbf{q}} = \lambda(\mathbf{q}) \mathbf{J}_{\mathbf{q}}$) if \mathbf{a} is either parallel or orthogonal to \mathbf{q} . The parallel eigenvalues are

$$\begin{aligned}
\lambda_{\parallel}(\mathbf{q}) &= \frac{i}{c} \left(k - \frac{|\mathbf{q}|^2}{k} \right) \hat{G}(\mathbf{q}) + Z \\
&= \frac{2\pi}{c} \left(1 - \frac{|\mathbf{q}|^2}{k^2} \right)^{\frac{1}{2}*} + Z
\end{aligned} \tag{23}$$

and the orthogonal eigenvalues are

$$\begin{aligned}
\lambda_{\perp}(\mathbf{q}) &= \frac{ik}{c} \hat{G}(\mathbf{q}) + Z \\
&= \frac{2\pi}{c} \left(1 - \frac{|\mathbf{q}|^2}{k^2} \right)^{-\frac{1}{2}*} + Z .
\end{aligned} \tag{24}$$

These eigenvalues are plotted in Figure 5. Note that the orthogonal currents have no charge accumulation, so λ_{\perp} includes only inductive and resistive effects. This is reflected in the fact that $\lambda_{\perp}(\mathbf{q}) \propto k$ for $|\mathbf{q}| \gg k$ (short range interactions), which is conform the factor ω in the inductive term in the EFIE (1). The longitudinal wave does have charge accumulation and therefore λ_{\parallel} includes capacitive effects. As a result we see that $\lambda_{\parallel}(\mathbf{q}) \propto 1/k$ for $|\mathbf{q}| \gg k$, which is conform the factor $1/\omega$ in the capacitive term in the EFIE (1).

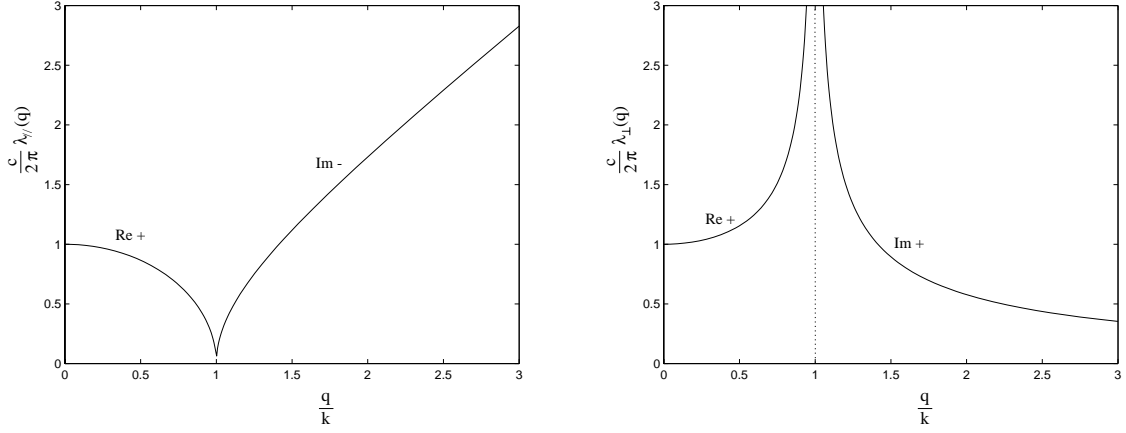


Figure 5: The continuous eigenvalues of the longitudinal ($\lambda_{||}$) and transversal (λ_{\perp}) current waves in absence of resistance. To include resistance Z should be added to the eigenvalue. The absolute value is shown while the text shows on which complex axis the values should be.

Note that, in absence of resistance, $\lambda_{||}(|\mathbf{q}| = k) = 0$ which means that there is no external field necessary to sustain this current, because the EFIE (1) is already satisfied. This also means that if there is a driving force for this mode, it will grow without restriction, which is called resonance. In a realistic problem, we have a (possibly small) resistance and we have a finite system size, but we can still expect damped resonances. In this case, a small driving force will result in large currents. This also means that the discretised system has a very small eigenvalue, which can slow down the convergence of an iterative solver.

We can also look at the eigenvalues and -functions of the electrostatic potential operator (9), corresponding to the matrix D introduced in section 4. Assuming a plane wave for the charge density $\rho_{\mathbf{q}}(\mathbf{x}) = e^{i\mathbf{q}\cdot\mathbf{x}}$, we get

$$\begin{aligned} V_{\mathbf{q}}(\mathbf{x}) &= \int G(\mathbf{x}' - \mathbf{x}) e^{i\mathbf{q}\cdot\mathbf{x}'} d^2x' \\ &= e^{i\mathbf{q}\cdot\mathbf{x}} \hat{G}(\mathbf{q}) , \end{aligned} \quad (25)$$

which shows that $\rho_{\mathbf{q}}(\mathbf{x})$ is an eigenvector with eigenvalue

$$\lambda_{\rho}(\mathbf{q}) = \hat{G}(\mathbf{q}) = -\frac{2\pi i}{k} \left(1 - \frac{|\mathbf{q}|^2}{k^2}\right)^{-\frac{1}{2}*} . \quad (26)$$

Since the D operator uses $\nabla \cdot \mathbf{J}$ as input instead of ρ , we get an extra factor i/ω according to equation (3), resulting in

$$\lambda_D(\mathbf{q}) = \frac{i}{\omega} \lambda_{\rho}(\mathbf{q}) = \frac{2\pi}{ck^2} \left(1 - \frac{|\mathbf{q}|^2}{k^2}\right)^{-\frac{1}{2}*} . \quad (27)$$

This has the same behaviour as $\lambda_{\perp}(\mathbf{q})$ in Figure 5b but with a k^{-2} scaling. We also see that $\lambda_D(\mathbf{q}) \propto 1/k$ for $|\mathbf{q}| \gg k$, as it should be.

A.2 The infinite wire

We will try to do the same for an infinite wire, as we did above for the infinite board. We will consider an infinitely long straight wire with radius R and surface impedance Z . Under the model

restriction that the current is in the tangential direction (parallel to the wire axis) and uniform around the wire, we have Fourier basis for the surface current consisting of

$$\mathbf{J}_q(\mathbf{x}) = \hat{\mathbf{z}} e^{iq\mathbf{x}_z} , \quad (28)$$

where $\hat{\mathbf{z}}$ is the unit vector in the tangential direction and \mathbf{x}_z is the tangential component of \mathbf{x} : $\mathbf{x}_z = \hat{\mathbf{z}} \cdot \mathbf{x}$. The resulting inductive field is given by

$$\begin{aligned} \mathbf{E}_q^L(\mathbf{x}) &= \frac{ik}{c} \int_{\Gamma} G(\mathbf{x}' - \mathbf{x}) \mathbf{J}_q(\mathbf{x}') d^2 x' \\ &= \frac{ik\hat{\mathbf{z}}}{c} \int_{\Gamma} G(\mathbf{x}' - \mathbf{x}) e^{iq\mathbf{x}'_z} d^2 x' \\ &= \frac{ik\hat{\mathbf{z}}}{c} e^{iq\mathbf{x}_z} \int_{\Gamma} G(\mathbf{y}) e^{iq\mathbf{y}_z} d^2 y \\ &= \frac{ik\hat{\mathbf{z}}}{c} e^{iq\mathbf{x}_z} \tilde{G}(q) , \end{aligned} \quad (29)$$

where we now introduced a special 1-dimensional Fourier transform of G on the wire surface Γ , given by

$$\begin{aligned} \tilde{G}(q) &= \int_{\Gamma} G(\mathbf{y}) e^{iq\mathbf{y}_z} d^2 y \\ &= \int_{\Gamma} \frac{e^{-ik|\mathbf{y}|}}{|\mathbf{y}|} e^{iq\mathbf{y}_z} d^2 y \\ &= 2R \int_{z=0}^{\infty} \int_{\theta=0}^{2\pi} \frac{e^{-ikR\sqrt{2-2\cos\theta+z^2}}}{\sqrt{2-2\cos\theta+z^2}} \cos(qRz) d\theta dz . \end{aligned} \quad (30)$$

Unfortunately, this can not be calculated analytically. Still, we continue with the capacitive field

$$\begin{aligned} \mathbf{E}_q^C(\mathbf{x}) &= \frac{i}{ck} \nabla \int_{\Gamma} G(\mathbf{x}' - \mathbf{x}) \nabla' \cdot \mathbf{J}_q(\mathbf{x}') d^2 x' \\ &= -\frac{iq^2}{ck} \hat{\mathbf{z}} e^{iq\mathbf{x}_z} \tilde{G}(q) \end{aligned} \quad (31)$$

and the resistive part

$$\mathbf{E}_q^R(\mathbf{x}) = Z \hat{\mathbf{z}} e^{iq\mathbf{x}_z} . \quad (32)$$

Adding this all together we find

$$\mathbf{E}_q(\mathbf{x}) = \left(\frac{ik}{c} \tilde{G}(q) - \frac{iq^2}{ck} \tilde{G}(q) + Z \right) \mathbf{J}_q(\mathbf{x}) , \quad (33)$$

which shows that \mathbf{J}_q is an eigenfunction with eigenvalue

$$\lambda(q) = \frac{ik}{c} \left(1 - \frac{q^2}{k^2} \right) \tilde{G}(q) + Z . \quad (34)$$

From this, we can already see that there is a resonance for $q = k$, unless a singularity in $\tilde{G}(q)$ changes this. This is not the case, since we know that this resonant mode is the so called transverse electromagnetic (TEM) mode [5, chapter 8]. Other resonant modes are not allowed by the model restrictions mentioned at the beginning of this section, and could only occur at very high frequencies ($\lambda \lesssim R$).

Just as we did for the infinite board, we can also look for the eigenfunctions and -values of the operator D for the wire. In exactly the same way, we get the eigenfunctions

$$\nabla \cdot \mathbf{J} = e^{iq\mathbf{x}_z} , \quad (35)$$

with eigenvalues

$$\lambda_D(q) = \frac{i}{\omega} \tilde{G}(q) . \quad (36)$$

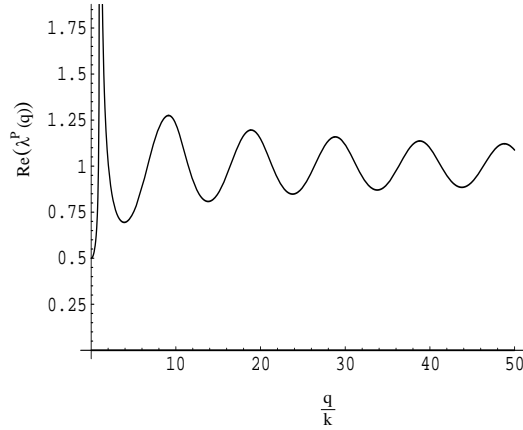


Figure 6: The continuous eigenvalues of the preconditioned operators ($\lambda^P(\mathbf{q})$) in absence of resistance, using a truncation distance of $R_T = 0.1\lambda$. Only the real part is shown, the imaginary part is small.

A.3 Preconditioning with truncated interaction

As described in section 5, we sparsify the matrix before we make the SAI preconditioner. We do this by throwing away all elements except those that correspond to interactions between physically close elements. We still expect good results for the highly oscillating eigenfunctions and we can verify this using Fourier analysis. The sparse matrix corresponds to a truncated interaction operator. We consider the operator truncated at distance R_T . For this operator all analysis above will hold, but now using the truncated Green function

$$G_T(\mathbf{y}) = \begin{cases} G(\mathbf{y}) & \text{for } |\mathbf{y}| \leq R_T \\ 0 & \text{for } |\mathbf{y}| > R_T \end{cases} . \quad (37)$$

This means that the eigenfunctions of the truncated operator are the same as above, but the eigenvalues are different, which implies that the full operator, preconditioned with the inverse of the truncated operator, also has the same eigenfunctions. The eigenvalues are those of the full operator divided by those of the truncated operator, leading to (for $Z=0$)

$$\lambda^P(\mathbf{q}) = \frac{\lambda(\mathbf{q})}{\lambda^T(\mathbf{q})} = \frac{\hat{G}(\mathbf{q})}{\hat{G}_T(\mathbf{q})} , \quad (38)$$

for the parallel, orthogonal and the “ D ” eigenvalues on the plane. The truncated Fourier transform \hat{G}_T can not be found analytically, but we can make a numerical approximation. In Figure 6 we show $\lambda^P(\mathbf{q})$ for an example where $R_T = 0.1\lambda$. We see that the high frequency tails of Figure 5 have been preconditioned quite well, but that the low frequency part still has a pole.

In practice, the interaction is not truncated outside a pure circle, but outside some discretised shape of the size of a few elements. Furthermore, we use an approximate inverse of this truncated operator, instead of the exact inverse. This means that, even in the limit of a very large board and very fine grid, we can expect deviations from this theoretical behaviour, but the general idea will still apply.

References

- [1] D. Colton and R. Kress, ‘Integral equation methods in scattering theory’, John Wiley & Sons, New York (1983).

- [2] S. C. Eisenstat, H. C. Elman and M. H. Schultz, *Variational iterative methods for nonsymmetric systems of linear equations*, SIAM J. Numer. Anal. **20** (2), 1983, pp. 345–356.
- [3] M.J. Grote and T. Huckle, *Parallel Preconditioning with Sparse Approximate Inverses*, SIAM J. Sci. Comp. **18** (3), 1997, pp. 838–853.
- [4] W. Hackbusch, ‘Multi-grid methods and applications’, Springer Verlag, Berlin (1985).
- [5] J.D. Jackson, ‘Classical Electrodynamics’, John Wiley & Sons, New York (1975).
- [6] Z. Q. Peng and A. G. Tijhuis, *Transient Scattering by a Lossy Dielectric Cylinder: Marching-on-in-Frequency Approach*, Journal of Electromagnetic Waves and Applications **7**, 1993, pp. 739–763.
- [7] Y. Saad and M.H. Schultz, *GMRES: a generalized minimal residual algorithm for solving nonsymmetric linear systems*, SIAM J. Sci. Statist. Comput. **7**, 1986, pp. 856–869.
- [8] P.P. Silvester and R.L. Ferrari, ‘Finite elements for electrical engineers’, Cambridge University Press, Cambridge, 1990
- [9] V. Simoncini and E. Gallopoulos, *An iterative method for nonsymmetric systems with multiple right-hand sides*, SIAM J. Sci. Comput. **16** (4), 1995, pp. 917–933.
- [10] H. A. van der Vorst and C. Vuik, *GMRESR: a Family of Nested GMRES Methods*, Num. Lin. Alg. with Appl. **1** (4), 1994, pp. 369–386.
- [11] C. Vuik, *Fast iterative solvers for the discretized incompressible Navier-Stokes equations*, Int. J. Num. Methods Fluids **22**, 1996, pp. 195–210.